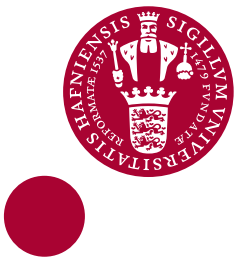


Temperature Dependence of Intrinsically Disordered Proteins in Simulations What are We Missing?

Jephthah, S.; Staby, L.; Kragelund, B. B.; Skepö, M.

Published in:
Journal of Chemical Theory and Computation

DOI:
10.1021/acs.jctc.8b01281

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2019, 15, 2672−2683

pubs.acs.org/JCTC

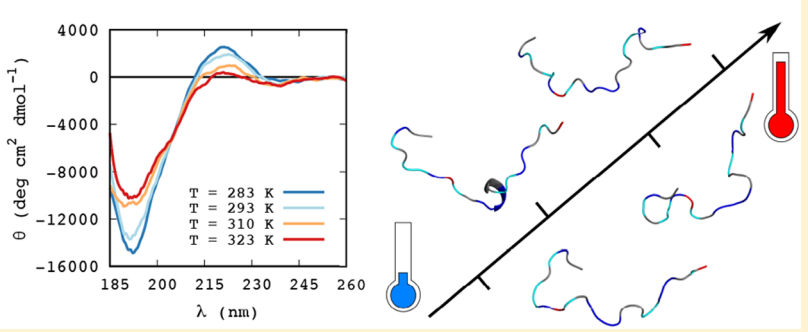
Temperature Dependence of Intrinsically Disordered Proteins in Simulations: What are We Missing?

S. Jephthah,*[,†] L. Staby,[‡] B. B. Kragelund,[‡] and M. Skepö*[,†]

[†]Division of Theoretical Chemistry, Lund University, 221 00 Lund, Sweden

[‡]Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, DK-2200 Copenhagen, Denmark

Supporting Information



ABSTRACT: The temperature dependence of the conformational properties in simulations of the intrinsically disordered model protein histatin 5 has been investigated using different combinations of force fields, water models, and atomistic and coarse-grained methods. The results have been compared to experimental data obtained from NMR, SAXS, and CD experiments to assess the accuracy and validity of the simulations. The results showed that neither simulations completely agreed with the experimental data, nor did they agree with each other. It was however possible to conclude that the observed conformational changes upon variations in temperature were not at all driven by electrostatic interactions. The final conclusion was that none of the simulations that were investigated in this study was able to accurately capture the temperature induced conformational changes of our model IDP.

1. INTRODUCTION

To understand the structural properties of intrinsically disordered proteins (IDPs), it is also necessary to study the possible conformational changes that could arise when the proteins are subjected to changes in temperature. While "'regular'" globular proteins are generally unfolded upon heating, IDPs have been reported to display the opposite behavior, showing temperature-induced partial folding or secondary structure formation instead.[1,2] Several studies have suggested that disordered proteins are in fact locally ordered, possessing a well-defined polyproline II (PPII) backbone structure, and that it is the destabilization of this structure that causes the contraction of IDPs at higher temperatures.[3−5]

Using computer simulations accompanied by experiments is a good way of studying conformational changes, as the simulations can give more detailed information about the system, while the experiments can validate the simulations as well as give an overview and an overall appreciation of the results. Several studies like this have been performed on IDPs, and quite a few of them have obtained good accuracy when it comes to comparing simulated and experimental results at room temperature. The results are unfortunately not as comparable when performing similar studies at other temperatures than room temperature. In the study by Wuttke et al.,[2] Monte Carlo simulations were performed with the OPLS/ABSINTH force field and implicit solvent, in which the force field was modified to also consider the temperature-dependent dielectric constant and solvation free energies. As a result of this, the temperature-induced collapse of a selection of disordered proteins was captured, which highlighted the importance of solvation effects in simulations of IDPs using implicit solvent. Temperature effects have however not been as easily replicated in molecular dynamics simulations with explicit solvent. In 2009, Nettels et al.[6] showed that the temperature dependence in protein simulations strongly depends on the selected force field and water model. They noticed that the AMBER ff03* force field[7] in combination with the TIP3P water model[8] gave an increase in the radius of gyration, R_g, with increased temperature, whereas a decrease in the R_g was found when the same force field was used in combination with the TIP4P-Ew water model,[9] which again stressed the importance of using an appropriate water model. Using the protein force field OPLS-AA/L with TIP3P also gave a slight collapse of the R_g with temperature.[10] Similar

Received: December 20, 2018
Published: March 13, 2019

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672−2683

results were observed several years later by Zerze et al.,[11] when protein simulations using the AMBER ff03ws force field[12] and the TIP4P/2005 water model[13] managed to qualitatively capture the temperature dependent collapse, although the Rg was always smaller than what experiments showed.

The fact that different water models give different results has been known for a long time. A small temperature dependence study (in which only two temperatures were tested) was done using the TIP4P water model[8] by Martin Neumann in 1986.[14] He showed that the dielectric constant decreases with temperature as it should, although the actual values were found to be too low. In 2004, Horn et al.[9] showed that the general tendency of decreasing dielectric constant with increased temperature was similar to the experimental values when using the TIP4P-Ew water model. The actual values were again too low, even though they were found to be considerably better than those obtained by using the TIP4P water model. Soon after, Abascal and Vega[13] determined that the dielectric constant at 298 K had better agreement with experiments when using the TIP4P-Ew water model compared to the TIP4P/2005 water model. Several studies have reported the dielectric constant of the TIP3P water model at room temperature to be higher than the expected value.[15−18] However, finding any studies in the literature concerning the temperature dependence of the dielectric constant of the TIP3P water model has been futile.

From the discussion above, it can be concluded that there have been several different studies concerning temperature dependence in simulations of proteins using different force field and water model combinations. While several simulations have shown to be sufficiently accurate at room temperature, the results from different temperatures were more varied. Thus, the question remains: *Is it possible for the currently available simulation methods to accurately mimic the experimental temperature induced structural changes in IDPs?*

In this case study, the temperature effects in simulations of the intrinsically disordered model protein histatin 5 (Hst5) have been investigated using both atomistic molecular dynamics (MD) and coarse-grained Monte Carlo (CG MC) simulations. Two of the currently available force fields appropriate for studying IDPs were chosen for the MD simulations: (i) AMBER ff99SB-ILDN[19] and (ii) CHARMM36m.[20] The AMBER ff99SB-ILDN force field was used in combination with the TIP4P-D water model,[21] but also the TIP3P water model,[8] which is known for causing too compact IDP structures when used in combination with this particular force field. The CHARMM36m force field was used in combination with the recommended TIP3P water model. For the CG MC simulations, we chose to use the Molsim simulation package (version 4.8.8).[22] To validate the simulation methods, the results were compared to data from various experimental methods such as small-angle X-ray scattering (SAXS), nuclear magnetic resonance spectroscopy (NMR), and circular dichroism spectroscopy (CD). The model protein, Hst5, is a small IDP that can normally be found in saliva, where it acts as a defense against fungal infections caused by, for example, *C. albicans*.[23,24] Hst5 is comprised of only 24 amino acid residues, of which almost 30% are histidines. In addition, seven of the other amino acid residues are positively charged, but only two are negatively charged. From this amino acid distribution, the conformational preference of Hst5 is predicted to be a polyampholytic coil or a polyelectrolytic semiflexible rod or coil, depending on the charge of the histidines.[25] When the histidines are charged, the protein obtains a more polyelectrolytic character. For the sake of simplicity, all simulated histidines were assumed to be neutral in this study. Illustrations of the atomistic and coarse-grained models of Hst5 are depicted in Figure 1.

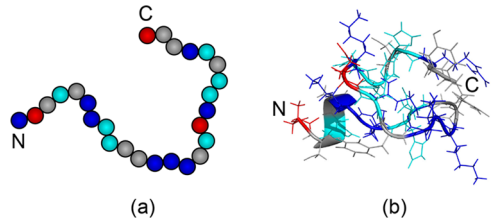


Figure 1. Illustration of the simulation models of Hst5. (a) The coarse-grained model and (b) the atomistic model overlaid with a cartoon representation to elucidate the backbone. The color scheme is as follows: red spheres represent the negative residues, blue spheres represent the positive residues, and the cyan spheres represent the histidine residues. The coarse-grained model also includes the protein termini as charged spheres.

2. METHODS

2.1. Computational Models and Methods. The model IDP, Hst5, has been simulated at different temperatures using several different combinations of simulation models and methods, see Tables 1 and 2. MC simulations were also performed for two additional CG peptides of equal length as Hst5 but with either (i) all residues being neutral or (ii) all residues having a charge of +1, see Table 2. Although all of the systems were included in the study, the main focus was on the MD simulations and the MC(IW) simulations (IW = implicit water).

Table 1. Abbreviations of the Different Combinations of Force Fields and Water Models Used in the All-Atom MD Simulations

abbreviation	force field	water model
MD(A3)	AMBER ff99SB-ILDN[19]	TIP3P[8]
MD(A4)	AMBER ff99SB-ILDN[19]	TIP4P-D[21]
MD(C3)	CHARMM36m[20]	TIP3P[8]

Table 2. Abbreviations of the Different Combinations of Chain Composition and Ionic Strength, *I*, Used in the CG MC Simulations

abbreviation	chain composition	*I* (mM)
MC_HST_10	Hst5	10
MC_NEU_10	all neutral	10
MC_POS_10	all charged	10
MC(IW)	Hst5	140
MC_NEU_140	all neutral	140
MC_POS_140	all charged	140

2.1.1. Atomistic MD Simulations. Molecular dynamics simulations were performed using the GROMACS package (version 4.6.7),[26−29] with the AMBER ff99SB-ILDN[19] and the CHARMM36m[20] force fields, together with the TIP3P[8] and the TIP4P-D[21] water models. The simulations were performed for two different kinds of simulation systems: (1) pure water and (2) Hst5. The water molecules were put in cubic boxes with edges of 2.5, 4.0, 5.0, 6.5, and 8.0 nm. A rhombic

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672−2683

dodecahedron was used as the simulation box for Hst5, with a minimum distance of 1 nm between the solute and the box edges. The protein systems were neutralized by the addition of five chloride ions. Periodic boundary conditions were used in all directions. The equations of motion were integrated using the Verlet leapfrog algorithm[30] with a time step of 2 ps. Energies and coordinates were saved every 5 ps. A Verlet list cutoff scheme was used for the nonbonded interactions, of which the short-ranged interactions were calculated using a pair list with a cutoff of 1 nm. Long-ranged dispersion interactions were applied to the systems' energy and pressure. The long-ranged electrostatics were managed by using Particle Mesh Ewald[31] with cubic interpolation and a grid spacing of 0.16 nm. A velocity-rescaling thermostat[32] and a Parrinello–Rahman pressure coupling[33] were used to keep the temperature and the pressure constant throughout the simulations. The simulations were performed using a pressure of 1 bar. Selected temperatures in the interval of 280–323 K were used for both the pure water simulations and the Hst5 simulations. All bond lengths were constrained using the LINCS algorithm.[34] The initial Hst5 model was built as a linear structure using PyMOL.[35] Initiations were performed in two steps: (1) 500 ps NVT simulations and (2) 1000 ps NPT simulations to stabilize the temperature and the pressure of the systems. The final production runs were simulated for 50 ns each for the water simulations, and for 5 × 1 μs each for the Hst5 simulations, with the exception of the MD(A4) simulation at 293 K, which was run for a total of 7 μs. No salt was used in any of these simulations since it has already been established in our previous studies that the addition of salt does not affect the structural properties of small peptides such as Hst5.[36,37]

2.1.2. Coarse-Grained Simulation Model. A coarse-grained model was used for the Monte Carlo simulations, in which each amino acid residue and the N- and C-termini are represented by hard spheres (monomers/beads) connected to each other by harmonic bonds (see Figure 1). The bead radius was set to 2 Å and the equilibrium bond length was set to 4.1 Å, with a force constant of 0.4 N/m. Each sphere was either neutral or charged according to the amino acid sequence at pH 7. The interactions between the nonbonded beads were simulated using a short-ranged attraction of 0.6 × 10^4 kJ Å^6/mol (corresponding to 0.6 kT), and an extended Debye–Hückel potential using the inverse Debye screening length, κ, and the dielectric constant, ε_r, of the solvent as input variables. Further details of the simulation model are described in the recent paper by Cragnell et al.[38]

2.1.3. Monte Carlo Simulations. The coarse-grained Metropolis Monte Carlo simulations were performed using an in-house extended version of the Molsim simulation package (version 4.8.8)[22] in the canonical ensemble (NVT). The coarse-grained Hst5 peptide was simulated in a cubic box with edges of 250 Å. Periodic boundary conditions were applied in all directions. The solvent in the box was implicit water with an ionic strength of 10 mM or 140 mM. The investigated temperatures in the simulation box were in the interval of 283–323 K. Each simulation was equilibrated for 10 × 2 ×10^5 steps, followed by 10 × 10^6 steps in the production run. Four different simulation moves were used throughout the simulations: (i) single bead translation, (ii) pivot rotation, (iii) chain translation, and (iv) chain slithering. Results from the MC(IW) simulations were used for the comparison with the MD simulations and the experiments. This particular system was chosen to facilitate the comparison with the experimental results in particular, in which physiological ionic strength was used.

2.1.4. Simulation Analyses. The average radius of gyration, $\langle R_g \rangle$ in the MD simulations was obtained by using the GROMACS tools g_polystat, and the dielectric constant was computed using g_dipoles. Frames for representative structures were obtained by using g_cluster and illustrated by using PyMOL.[35] The minimum distance between the periodic images of each MD simulation was calculated using g_mindist. By monitoring that the minimum distance never approached the cutoff distance of the nonbonded interactions in the system, it was ensured that the simulated protein chain did not interact with its periodic images. Theoretical SAXS intensities from the MD simulations were obtained using CRYSOL (version 2.8.2). The secondary structure was analyzed using the DSSP program (version 2.2.1),[39] with modifications according to Chebrek et al.[40] to include analysis of the PPII structure. Probability density functions of the R_g, $p(R_g)$, were used to analyze and confirm the accuracy of the sampling in the simulations (see the Supporting Information). Estimations of the full width half-maximum, FWHM, of the $p(R_g)$ were obtained according to Cragnell et al.,[38] that is, by fitting the $p(R_g)$ curves with the Gaussian function

$$f(x) = a \cdot \exp\left[-\frac{(x-b)^2}{c^2}\right] \quad (1)$$

and computing the FWHM according to the following equation:

$$\mathrm{FWHM} = 2c\sqrt{\ln(2)} \quad (2)$$

The variables a, b, and c are fitting parameters, and the FWHM values were reported with a 95% confidence interval.

2.2. Experiments and Techniques. Three experimental techniques have been used in this study: (i) SAXS, (ii) NMR, and (iii) CD. Two buffers were used in the experiments: (i) a 20 mM Na_2HPO_4/NaH_2PO_4 phosphate buffer at pH 7.0 and (ii) a 20 mM Tris buffer at pH 7.5. The phosphate buffer would be preferred to use in all cases due to its ideal buffer range and negligible temperature dependence. Unfortunately, the phosphate buffer was not suitable for the SAXS experiments, in which it has been shown to cause severe radiation damage. Thus, the Tris buffer was used for these experiments instead but with a slightly increased pH to maintain a decent buffering action. The ionic strength of the solutions was set to 140–150 mM in the SAXS and CD experiments. This was done to mimic physiological conditions and to exclude any intermolecular interactions. The effect of the ionic strength is, however, assumed to be negligible because of the small size of the Hst5 peptide, which has also been seen in previous SAXS measurements and simulations.[36,37] Thus, the NMR measurements remained salt-free in this study.

2.2.1. SAXS and CD Sample Preparations. The Hst5 peptide was obtained from Genemed Synthesis Inc. (San Antonio, TX, USA) as white powder with 95.73% purity and trifluoroacetate (TFA) as counterion. Tris buffers (Saveen Werner AB, >99.9% purity, CAS Registry no. 77-86-1) for the SAXS measurements were prepared at a concentration of 20 mM in Milli-Q water, and acidified with HCl so that the buffers would maintain pH 7.5 at the appropriate measuring

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672–2683

temperatures. Calculations of the pH at different temperatures were performed by using a $d(pK_a)/dT$ value of −0.028. Sodium phosphate monobasic monohydrate (NaH_2PO_4, Sigma-Aldrich, ≥99.0% purity, CAS Registry no. 10049-21-5) and sodium phosphate dibasic dihydrate (Na_2HPO_4, Sigma-Aldrich, ≥99.5% purity, CAS Registry no. 10028-24-7) were used to prepare a 20 mM phosphate buffer in Milli-Q water for the CD measurements. The pH of the phosphate buffer was set to 7.0 with NaOH. The ionic strength of the buffer solutions was set to 140−150 mM. NaCl was used to adjust the ionic strength for the SAXS measurements, and NaF (VWR International, ≥99.5% purity, CAS Registry no. 7681-49-4) was used for the CD measurements. Before any peptide was added, the buffers were filtered through a 0.2 μm hydrophilic polypropylene membrane (Pall Corporation). After the peptide was dissolved, a concentration cell (Vivaspin 20, 2.0 kDa MWCO, product no. VS02H92, Sartorius Stedim Biotech GmbH, Göttingen, Germany) was used to remove low molecular weight impurities. The samples were rinsed with buffer by centrifugation to a maximum of 3500 RCF at 15−18 °C, and the amount of buffer used for rinsing was equal to at least 10 times the sample volume. Dialysis of the SAXS samples was done using dialysis cups (Slide-A-Lyzer MINI, 2.0 kDa MWCO, product no. 69580, Thermo Scientific, United States), and the CD samples were dialyzed using dialysis cassettes (Slide-A-Lyzer Dialysis Cassettes, 2.0 kDa MWCO, product no. 66205, Thermo Scientific, United States). All samples were dialyzed for at least 12 h in a buffer volume of at least 100× the sample volume to ensure exact background.

2.2.2. SAXS Measurements. SAXS measurements were performed at beamline BM29 at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. The incident-beam wavelength was 0.99 Å, and the distance between the sample and the PILATUS 1 M detector was 2867 nm, giving a scattering vector range of 0.037−4.928 nm^{-1}. For each sample and pure solvent, at least ten successive 1 s frames were recorded and analyzed. Special attention was paid to radiation damage by comparing the successive frames prior to further processing of the data. The measurements were performed at selected temperatures of 10, 20, 37 and 50 °C. Before the SAXS measurements, the samples were further centrifuged at 18400 RCF and 6 °C for at least 30 min to remove aggregates. Protein concentrations were measured after preparation and again immediately before the SAXS measurements using a Nanodrop spectrophotometer (ε = 2560 M^{-1} cm^{-1}, λ = 280 nm). To obtain the proper SAXS spectra, the background (i.e., the pure solvent) was subtracted from the corresponding sample spectrum. Normalization was done by converting $I(0)$ to an absolute scale by measuring the scattering of pure water.

2.2.3. CD Measurements. Far-UV (185−260 nm) CD spectra were obtained using a Jasco J-715 CD spectrometer with a model PTC-348WI Peltier type temperature control system (Hachioji, Tokyo, Japan). The samples were filtered though a Millex-GV Filter with a pore size of 0.22 μm (Merck Millipore Ltd., Ireland), before being added to a quartz cuvette with a path length of 0.1 cm. Selected temperatures of 10, 20, 37, and 50 °C were used for the measurements. The scanning rate was 20 nm/min, with a bandwidth of 2.0 nm, and a response time of 2 s. At least five scans were performed and averaged for each measurement. All spectra were smoothed using Savitzky−Golay filtering (SGF) with a bandwidth of 75. The background spectra were subtracted from the respective sample spectra and checked for oversmoothing before presenting the final data (see example in the Supporting Information). The smoothing process was done following the protocol of Norma J Greenfield.[41]

2.2.4. NMR Measurements. All NMR experiments were acquired on a Bruker Avance III 600 MHz (^1H) spectrometer equipped with a TCI cryogenic probe. Samples contained 1 mM Hst5 in a 20 mM mixture of Na_2HPO_4/NaH_2PO_4 (pH 7.0), 10% D_2O, 0.25 mM DSS, and 0.25% (v/v) 1,4-dioxane. Pulsed-field gradient (PFG) NMR diffusion experiments were recorded using a standard Bruker PFG-LED pulse sequence with bipolar gradients and solvent presaturation during the relaxation delay of 3 s.[42] A total of 32 spectra with gradient strengths ranging from 2% to 98% of its maximum value were recorded. The pseudo-2D data was Fourier transformed and baseline corrected in Topspin (Bruker) and subsequently processed in Dynamics Center (Bruker) to obtain peak intensities. Diffusion constants were determined by fitting of peak intensity decays against the gradient strength using the Stejskal−Tanner equation:

$$I = I_0 e^{-g^2\gamma^2\delta^2(\Delta-\delta/3)D} \tag{3}$$

where I is the intensity, g the gradient strength, γ the gyromagnetic ratio of ^1H, δ the gradient length, Δ is the diffusion time, and D is the diffusion constant. The diffusion constants of a selection of peaks from the aromatic and aliphatic side-chain regions were plotted as histograms and fitted to a Gaussian

$$f(x) = A \cdot e^{-(x-\mu)^2/2\sigma^2} \tag{4}$$

where μ is the mean, σ the standard deviation, and A the amplitude. Different bin sizes where tested to check the robustness of the fit. A small amount of 1,4-dioxane with a known hydrodynamic radius, $R_{H,ref}$, of 2.12 Å was used as an internal reference and used for the calculation of $R_{H,prot}$ according to Wilkins et al.[43]

$$R_{H,prot} = \frac{D_{ref}}{D_{prot}} \cdot R_{H,ref} \tag{5}$$

3. RESULTS AND DISCUSSION

In this study, both atomistic MD simulations and coarse-grained MC simulations of Hst5 have been performed at different temperatures. Due to the extensive computer resources needed for the MD simulations, only four temperatures were tested with this method. By using a coarse-grained model for the MC simulations, a substantial amount of computational time was saved, which enabled the study of a wider range of temperatures. Analysis of the simulated data provided R_g probability density functions, as well as the average radius of gyration, $\langle R_g \rangle$, and SAXS profiles of each simulated system. Secondary structure analysis could also be obtained from the MD simulations. Discussion about convergence and sampling is referred to the Supporting Information. The results from the experimental methods, that is, SAXS, NMR, and CD, were utilized to assess the validity of the simulations and to discern any temperature-dependent conformational changes. In this section, we will start by presenting and discussing changes in $\langle R_g \rangle$ for all the different methods that have been used in this study (if applicable). Subsequently, the shape, flexibility, and secondary structure of

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672−2683

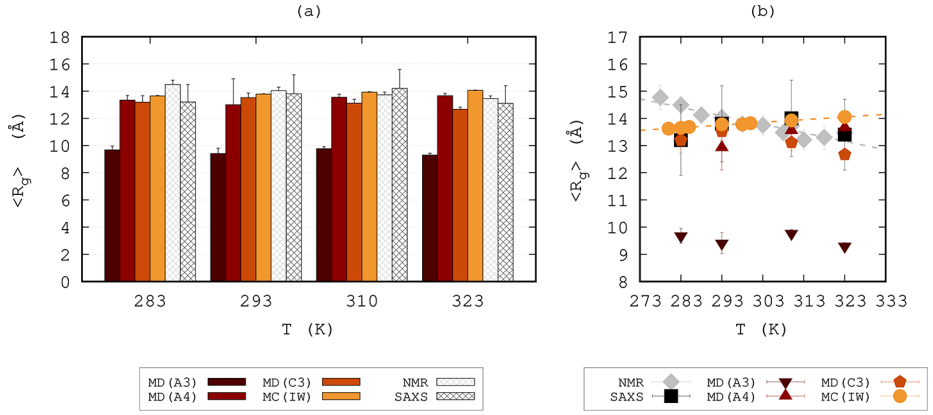


Figure 2. Radius of gyration, $\langle R_g \rangle$, as a function of temperature from all six methods. (a) Histograms to emphasize the difference in $\langle R_g \rangle$ between the methods at the different temperatures. (b) The change in temperature of the different methods. The dashed lines in panel b are linear regression trend lines. All data points include error bars, but the errors are too small to be visible in some cases.

Table 3. $\langle R_g \rangle$ of All Methods at the Four Selected Temperatures

	$\langle R_g \rangle$ (Å)					
T (K)	SAXS	NMR	MD(A3)	MD(A4)	MD(C3)	MC(IW)
283	13.2 ± 1.3	14.49 ± 0.31	9.68 ± 0.28	13.34 ± 0.33	13.19 ± 0.47	13.65 ± 0.03
293	13.8 ± 1.4	14.03 ± 0.26	9.41 ± 0.39	12.92 ± 0.82	13.52 ± 0.34	13.78 ± 0.03
310	14.0 ± 1.4	13.73 ± 0.20	9.77 ± 0.15	13.55 ± 0.22	13.11 ± 0.29	13.92 ± 0.04
323	13.4 ± 1.3	13.46 ± 0.19	9.30 ± 1.13	13.66 ± 0.18	12.67 ± 0.15	14.05 ± 0.03

our model peptide will be analyzed before the finalizing concluding remarks.

3.1. Temperature Correlations in the Conformational Ensemble. The $\langle R_g \rangle$ at different temperatures was obtained from both the MD and the MC simulations, as well as from SAXS and NMR. A conversion factor of 1.1 was used to estimate the $\langle R_g \rangle$ from the hydrodynamic radius, R_H, which was obtained from the NMR experiments. The conversion factor was computed using the method presented by Nygaard et al.[44] The $\langle R_g \rangle$ from SAXS was obtained from PRIMUS using the Guinier approximation with $qR_g(\max) < 0.8$. A comparison between the $\langle R_g \rangle$ from all different methods is presented in Figure 2. Figure 2a emphasizes the difference in $\langle R_g \rangle$ between all methods at the selected temperatures, and Figure 2b illustrates the $\langle R_g \rangle$ changes as a function of temperature. Linear regression analysis was done on all data sets, assuming statistical significance for $p > 0.05$ (see Table S1 for more detailed information). The $\langle R_g \rangle$ values at four selected temperatures are also displayed in Table 3. Visual inspection of Figure 2a revealed that the $\langle R_g \rangle$ from all methods except MD(A3) were similar at the selected temperatures. The reason for the lower $\langle R_g \rangle$ values with this force field and water type combination has already been frequently and thoroughly discussed[19,45−48] and will, thus, not be explained here. However, the observed trends in Figure 2b did give a different view of the results: while the NMR data displayed a distinct decreasing behavior with increased temperature, the $\langle R_g \rangle$ from the SAXS measurements showed a small increase until the peak value of 310 K, after which it decreased again. Taking the linear regression analysis into consideration, the $\langle R_g \rangle$ from the NMR data was confirmed to have a strong and significant correlation with the temperature. The same could not be concluded about the SAXS results, for which no linear correlation could be identified. The $\langle R_g \rangle$ from the MD(A3) and the MD(C3) simulations showed a correlation that was similar to the NMR results, although weaker and not significant in the selected confidence level. An opposite correlation between the $\langle R_g \rangle$ and the temperature was found in the MD(A4) simulations compared to the other MD simulations, but this correlation was also proven to be insignificant. The results from the MC(IW) showed a positive correlation with increasing temperature, which was also confirmed to be the only simulated correlation of statistical significance. Because of the known problem of sampling,[46] the lack of significant correlation in the MD simulations was not surprising. However, the change in $\langle R_g \rangle$ that would arise because of improved sampling would most likely not be of enough magnitude to change the observed correlations, based on the results found by Hicks and Zhou,[49] where they only observed small differences (≤0.5 Å) for the $\langle R_g \rangle$ with and without enhanced sampling methods. This value corresponds to approximately one-fourth of the bead radius in our coarse-grained model, and is only slightly larger than the error estimates of the MD simulations. In fact, the $\langle R_g \rangle$ only changed from 13.00 to 12.92 Å upon extending two of the replicates in the MD(A4) simulation from 1 to 2 μs.

Since the SAXS and NMR experiments gave different results, it is difficult to determine which one of the temperature-dependent conformational changes is the most reliable. There were several basic factors differing between the two experiments: (i) the buffer, (ii) the pH, and (iii) the ionic strength. However, as already mentioned, neither of these differences is expected to give rise to such huge deviations as was seen in the results. Thus, we are looking for answers elsewhere in future studies. One observation that was done in the SAXS experiments at 310 K and higher, was that the computed

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672−2683

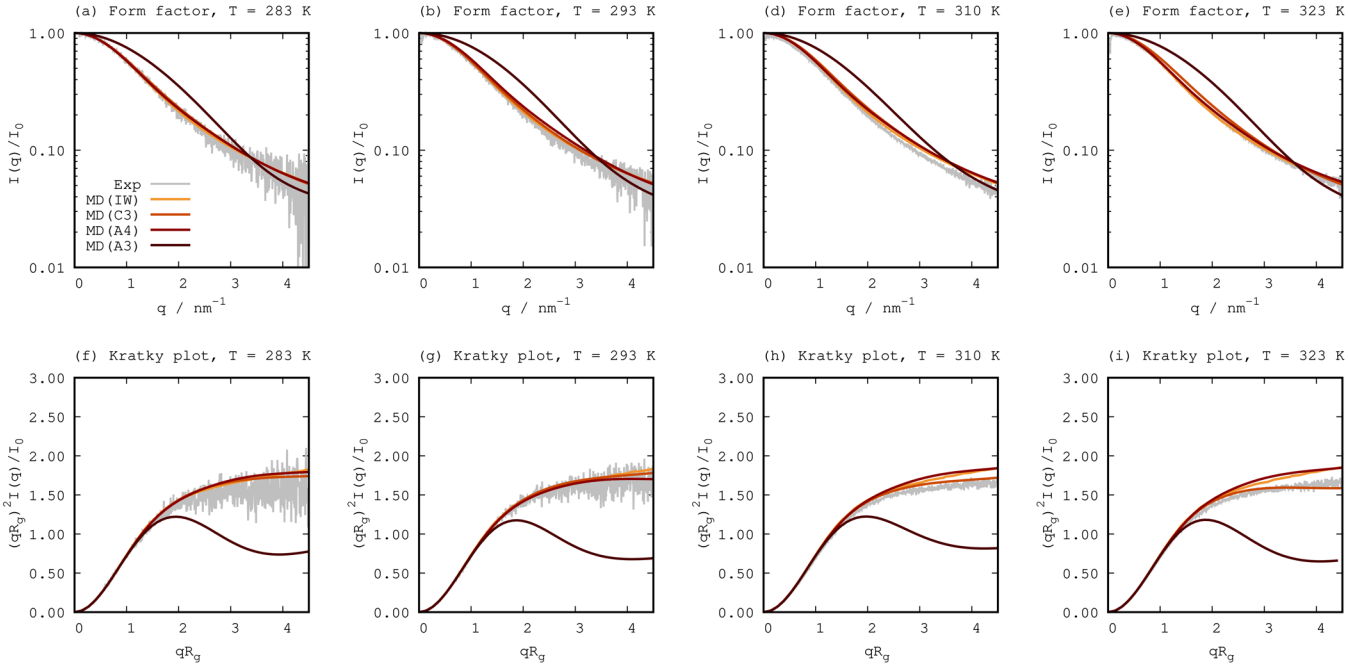


Figure 3. Form factor (upper row, a–e) and the normalized Kratky plot (lower row, f–i) from SAXS measurements (gray) compared to computed SAXS curves from simulations at four different temperatures (left to right).

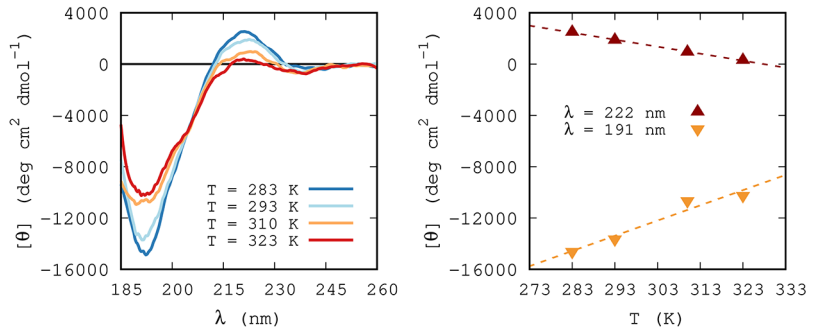


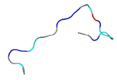
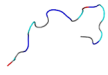
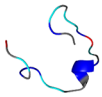
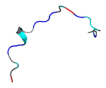
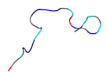
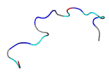
Figure 4. (a) Normalized CD spectra of Hst5 at different temperatures in samples with an ionic strength of 150 mM and (b) the change in ellipticity at wavelengths of 191 and 222 nm with temperature.

molecular mass became significantly lower (i.e., a deviation of more than 10%) for samples of low concentration, which could potentially mean that there is some sort of destabilization of the peptide under these conditions. This is however only a speculation, and the SAXS data presented in this paper had their molecular masses within the acceptable deviation limits. Another possibility is that the change in $\langle R_g \rangle$ is too small for being properly distinguished by the SAXS.

3.2. Shape and Flexibility. By studying the SAXS spectra, information on the shape and flexibility of the peptide can be obtained. Comparing the experimental data to that from simulations provide an important tool in validation of the simulation models. Comparison of the results from the simulations and the experiment in Figure 3 makes it clear that the MD(A4), the MD(C3), and the MC(IW) simulations managed to imitate the experimental form factor very well at 293 K (Figure 3b). The normalized Kratky plot at this temperature (Figure 3g) also agree on the shape of the peptide, which from the plateau indicated a high degree of flexibility. The MD(A3) simulations did not agree with the experiments, which can be seen by the indication of a more globular shape in both the form factor and the Kratky plot. All SAXS results at 293 K were in agreement with previous studies.[36,47] Visual inspection of the form factor suggested that the agreement between simulations and experiment seems to hold at 283 and 323 K but not at 310 K. Comparison of the Kratky plot does, however, show larger differences between the simulations and the experiments at all temperatures other than 293 K. The deviations are more pronounced at higher qR_g values, where the simulations generally have slightly higher values. It is difficult to discern if this difference is significant or not since the Krakty plot is very sensitive in this region. Even small alterations due to, for example, the background subtraction can cause the Kratky plot to change at high qR_g values. Steeper positive slopes usually indicates increasingly extended or rigid structures, but the difference is expected to be larger for such a significant structural change. Furthermore, it is important to keep in mind that experimental SAXS has quite low resolution compared to the SAXS curves from the simulations. Another reason for the differences between the simulations and the experiments could be that the histidines are assumed to be neutral throughout all of the simulations, even though histidines are known to charge regulate in solution. This

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672–2683

Table 4. Cartoon Depiction of the Representative Structures from the MD Simulations and Their Corresponding Secondary Structure Per Amino Acid According to DSSPPII Analysis

T (K)	MD(A3)	MD(A4)	MD(C3)	MD(A4)/MD(C3)
283	 -TTT-EETTEE--TTT-S--	 -STT----SS--TTS--	 --S-TT-SS-TT-S-S--	
293	 -TTTTS-SS--S-SS-TT--	 -TTS-TTS-S-SS-SS--	 ----STTTTT----S--	
310	 -TT-S-SGGGT-S--TTT-	 --SS-STT--S-S--TT--	 -S-SS--SS-SSTTTSGGG-	
323	 -HHHHTT-STT-SS--S---	 -SSS-TTSSS--SS-S--	 --SSSS---SSS-SS--	

does however seem unlikely since a previous study by Kurut et al.[54] has shown that the charge regulation of the histidines is very low at neutral pH in simulations. Additionally, protonation of the histidines results in a more extended structure due to the additional electrostatic repulsion that accompanies the protonation (see Figure S4). Inspection of the experimental Kratky curves proved no change in the shape and flexibility of the peptide with the temperature. The differences between the simulated SAXS spectra at different temperatures were also found to be visually negligible for each model and are thus only shown in Figure S5.

3.3. Secondary Structure. Temperature dependence in the secondary structure was studied by CD spectroscopy, see Figure 4a. The spectra showed a negative band around 191 and 240 nm, as well as a positive band around 222 nm. The absolute intensity of the bands was shown to decrease as the temperature was increased. The change in ellipticity at 191 and 222 nm as a function of temperature is illustrated in Figure 4b. Linear regression analysis revealed a significant correlation between the ellipticity and the temperature at the selected bands (see Table S1), which indicates destabilization of PPII structure with increasing temperature.[55] Since the PPII structure is quite extended, one could speculate that the peptide becomes less extended when the temperature is increased, which agrees very well with the NMR results. This is, however, not supported by the SAXS results, which indicated no significant change in shape and flexibility with the temperature. Shifts in the bands toward longer wavelengths were observed, which is also associated with a destabilization of the PPII structure. Measurements at high temperatures and shorter wavelengths are however less reliable due to increased tension (HT) voltage. This makes it difficult to deduce whether or not this observation at 191 nm was the expected natural response with some discrepancies because of the limitations of the method or if it was a random variation because of insufficient resolution.

Cluster analysis was performed to acquire a central, or representative, structure of each simulation, on which secondary structure analysis was performed in addition to the complete MD trajectories (see Figure S6). Contrary to the experiments, the representative structures did not show any traces of PPII structure for any force field or temperature, see Table 4. Instead the representative structures were found to be dominated by bends (S), turns (T), and random coils (−). These structures were also dominating in the complete simulation trajectories. As expected, the MD(A3) simulations gave rise to more compact turn and helical structures (H, α-helix; G, 3_{10}-helix) at a temperature of 293 K and above. At 283 K, the central part of the peptide was instead interpreted as an extended strand (E). The turn content was more distinguished in the MD(C3) simulations compared to the MD(A4) simulations. One deviation from the unstructured behavior was found for MD(C3) at 310 K, which representative structure suggested a 3_{10}-helix toward the C-terminus of the protein. Visual inspection of the representative structures from the MD(A4) and MD(C3) simulations showed similar conformations and are, thus, aligned in the last column of the table to visually enhance the comparison. In the simulations using the AMBER force fields, a few residue sequences were found to consistently adopt turn conformations. In the MD(A3) simulations, the preserved structure was found around His-3, Ala-4, and Lys-5. These residues were found in a turn conformation at all temperatures except 323 K, where they were found in an α-helix conformation instead. The turn in the MD(A4) was not present at 283 K but could otherwise be found at His-8 and Gly-9. The MD(C3) was not seen to possess such specific structure propensity. It is not unreasonable to expect turns around these type of amino acids since they are known to be disorder-promoting.[56] However, further studies are required to discern if this structural preservation is a true effect or simply due to insufficient sampling. While $\langle R_g \rangle$ might not change significantly, the

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672−2683

secondary structure analysis is more likely to benefit from enhanced sampling. Although improbable, inadequate sampling could also be a reason for the lack of PPII structure in the simulations. Only extending the simulations is not likely to be sufficient to change the conformation of the representative structure of the simulation. This was observed when comparing the 5 μs version of the MD(A4) simulation to the extended 7 μs version, for which the structures were identical (see Figure S7). However, since we are only interested in the overall secondary structure content and not specific details about every possible conformation, and although it would be interesting to expand the study using enhanced sampling methods, this is out of scope in our study.

3.4. Temperature-Dependent Interactions in the Simulations. To exclude the possibility that the water models could cause an inaccurate temperature dependence in the MD simulations, pure water simulations were also performed using the AMBER ff99SB-ILDN force field in combination with the two water models. The dielectric constant was computed from the water simulations at different temperatures, and compared to the reference values, see Figure 5. The figure showed that the dielectric constant of both the TIP3P and the TIP4P-D water model decreased with increased temperature, as expected. The actual magnitudes of the obtained dielectric constants were, however, shown to deviate, resulting in larger values than the reference values while using the TIP3P water model, and values smaller than the reference values while using the TIP4P-D water model. Since decreasing correlations were captured in the water models, that is, a correct temperature dependence was observed, any possible erroneous contribution to the behavior of the Hst5 simulations can most likely be disregarded.

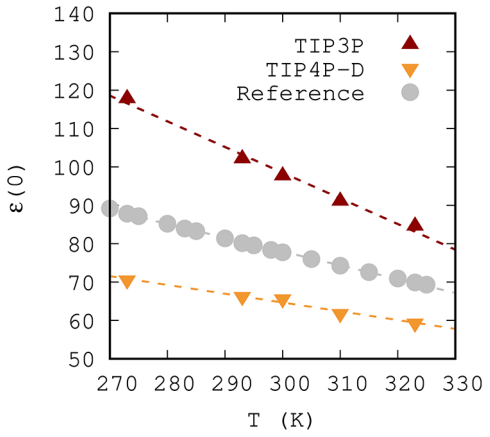


Figure 5. Dielectric constant of the TIP3P and the TIP4P-D water models at different temperatures compared to the reference values from Malmberg and Maryotte (1956),[50] Owen et al. (1961),[51] Uematsu and Frank (1980),[52] and Fernandez et al. (1997).[53]

Further investigation of the dielectric constant in the simulations was done by performing additional MC simulations using the dielectric constants from the MD simulations instead of the standard reference values. The resulting $\langle R_g \rangle$ values were found to be basically identical to those from the MC(IW) simulations (and are thus not shown), which indicates that the electrostatic contribution is not dominating the temperature dependence in the MC simulations. This observation was further explored by performing additional coarse-grained MC simulations of two other chains of equal length as Hst5, but consisting completely of (i) neutral beads and (ii) beads with a charge of +1 each (see Table 2). The results are shown in Figure 6, where the temperature dependence of the $\langle R_g \rangle$ was found to be the same regardless of what fraction of charge the chain possesses. The temperature dependence was also found to be independent of the ionic strength. These results combined strengthens the notion of that the temperature dependence is not electrostatically driven in the coarse-grained MC simulations.

The strength of the electrostatic interactions is dependent on the temperature through the dielectric constant, and the interactions become stronger at higher temperatures. Consequently, if the electrostatic interactions were dominating, the simulated peptide would most likely either (i) become more extended at higher temperatures because of increasing repulsion between the positive residues in the chain, (ii) show an increased interaction between residue Glu-16 and any of the positive amino acid residues, or (iii) display a combination of these two effects. Since neither of the proposed effects was observed in the MD results, this could indicate that the effect of temperature on the conformational properties of

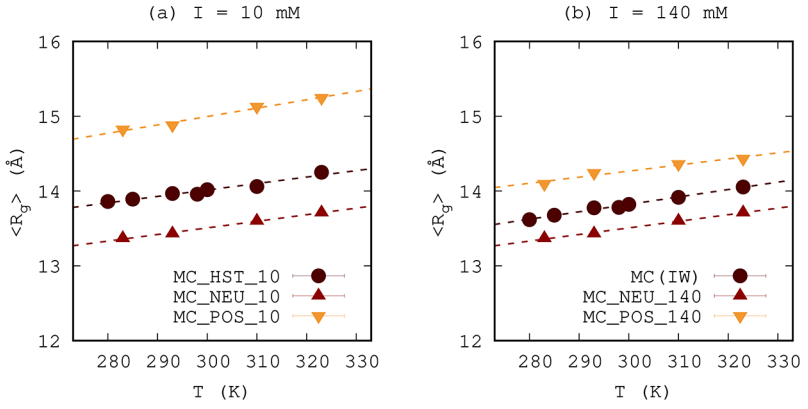


Figure 6. Comparison of the change in $\langle R_g \rangle$ with temperature for three chains of equal length but with different charge compositions. The simulations were performed using two different ionic strengths: (a) 10 and (b) 140 mM. See Table 2 for a more detailed description of the simulated systems.

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672−2683

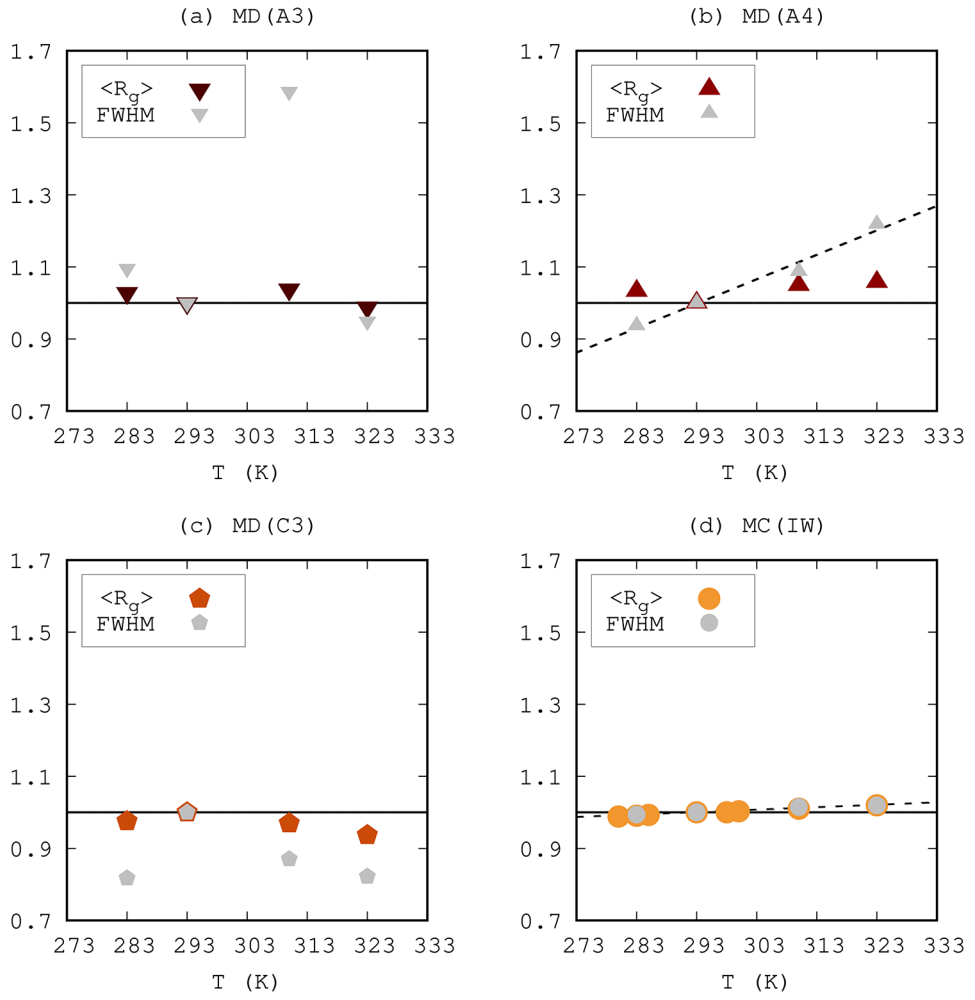


Figure 7. $\langle R_g \rangle$ compared to the FWHM as a function of temperature for the four different simulation methods. All points are normalized with their respective value at room temperature.

Hst5 is not primarily driven by the electrostatic interactions in the MD simulations.

Assuming that the electrostatic interactions are not the main driving force for the temperature dependence of conformational changes in Hst5 in the simulations leaves only one option: the entropic effect. To evaluate the conformational entropy in the simulated systems, the FWHM of $p(R_g)$ (Figure S2) was determined. The results were normalized and compared to the normalized $\langle R_g \rangle$, see Figure 7. Linear correlation between the FWHM and the temperature was found for the MD(A4) and the MC(IW) simulations only. Since the correlations were found to be positive, this means that the chain entropy was increasing with the temperature. Unfortunately, several of the data points of the remaining two methods cannot be considered reliable because of bad Gaussian fitting caused by seemingly two or more radius distributions under the same curve (e.g., 310 K in Figure S2a). Thus, another fitting model or improved sampling would be needed to properly evaluate the conformational entropy of the MD(A3) and MD(C3) simulations. Another observable correlation was identified between the $\langle R_g \rangle$ and the FWHM in the MC(IW) simulation, which suggests that the conformational changes heavily depend on the entropic forces. To investigate the correlation further, linear regression analysis was performed to prove its significance (see Figure S3 and Table S1).

To summarize, the results indicated that the temperature dependence of the conformational properties of Hst5 most likely is driven more by entropic effects than by the electrostatic interactions in the simulations, especially in the case of the MC simulations. This observation is in agreement with the loss off PPII structure that was observed in the experiments. Nevertheless, neither of the simulations was able to predict any accurate temperature dependence of the conformational ensemble, which gives rise to the question in the title of this study: What are we missing? Considering the MD simulations, both the TIP4P-D water model and the CHARMM36m force field have been modified to sustain more favorable protein−water dispersion interactions. As mentioned in the introduction, MC simulations with added solvation free energies have also been shown to give more accurate results. Thus, although several studies have shown that proper balancing of protein−solvent interactions is a key element to obtain good simulation results, it is obviously not enough to only scale these interactions to capture conformational properties of IDPs at other temperatures than room temperature. In addition, since the electrostatic interactions and the entropic effects are the only temperature dependent interactions in the simulations, these are either not completely

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672−2683

accurate, or it might also be needed to adjust the strength of the short-ranged interactions with respect to the temperature.

4. CONCLUSION

The case study presented here has focused on investigating the accuracy and validity of several simulation models and methods considering temperature induced structural changes. The simulated results have been compared to each other, but also to experimental results. As noted in several previous studies, we again concluded that the choice of force fields and water models heavily influences the outcome of the MD simulations, and also when it comes to temperature effects. It was observed that all three studied MD systems displayed different responses upon changes in temperature. Weak correlations between the temperature and the $\langle R_g \rangle$ were seen, but these were statistically insignificant, which is most likely attributed to the known problem of insufficient sampling. While enhanced sampling would decrease the errors and stabilize the values, and thereby also give more prominent correlations, it is unlikely that it would change the general trends observed in this study. The only simulation method that gave an actual significant temperature correlation was the CG MC method. This correlation was not supported by the NMR experiments, which instead suggested the opposite temperature dependence. Secondary structure analysis of the MD simulations was also unable to match the experimental results. The CD strongly indicated a loss of PPII structure, but this structure was not found in any of the simulated systems, and neither did these systems show any consistent correlation in secondary structure changes. However, since the secondary structure analysis was based on the representative structure from each simulation, enhanced sampling could possibly lead to a different distribution of secondary structure without significantly change averages of the other studied properties. Furthermore, we also deduced that it is unlikely that the water models are responsible for the different temperature behaviors and that the temperature induced correlations are not electrostatically driven in neither of the simulation methods. In conclusion, this study has shown that several of the simulation models and methods that are currently available for simulations of IDPs do *not* manage to capture the temperature induced structural changes, since none of the methods presented here was able to accurately explain the experimentally observed correlations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.8b01281.

Notes on CD subtraction and smoothing, as well as simulation convergence and sampling; linear regression statistics, additional FWHM plots, and supplementary Kratky plots; and time evolution of the secondary structure and supplementary representative structures from the MD simulations (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: stephanie.jephthah@teokem.lu.se. Phone: +46-46-222 17 54.

*E-mail: marie.skepo@teokem.lu.se. Phone: +46-46-222 33 66.

ORCID

S. Jephthah: 0000-0003-4287-506X

M. Skepö: 0000-0002-8639-9993

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge financial support from NordForsk's Neutron Science Programme, the Crafoord Foundation, and the Danish Research Council (#4181-00344 to BBK). Computer resources for the simulations were provided by the Swedish National Infrastructure for Computing (SNIC) at the Center for Scientific and Technical Computing at Lund University (LUNARC). We also thank the European Synchrotron Radiation Facility (ESRF), Grenoble, France, for providing beamtime, and Dr. Bart van Laer for assistance at beamline BM29. Finally, we would also like express our gratitude to Dr. João Henriques and Carolina Cragnell for providing valuable guidance regarding computer simulations and SAXS experiments, respectively.

■ REFERENCES

(1) Uversky, V. N. Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J.* **2009**, *28*, 305−325.

(2) Wuttke, R.; Hofmann, H.; Nettels, D.; Borgia, M. B.; Mittal, J.; Best, R. B.; Schuler, B. Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 5213−5218.

(3) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. Conformation of the backbone in unfolded proteins. *Chem. Rev.* **2006**, *106*, 1877−1897.

(4) Kjaergaard, M.; Nørholm, A.-B.; Hendus-Altenburger, R.; Pedersen, S. F.; Poulsen, F. M.; Kragelund, B. B. Temperature-dependent structural changes in intrinsically disordered proteins: Formation of α−helices or loss of polyproline II? *Protein Sci.* **2010**, *19*, 1555−1564.

(5) Sreerama, N.; Woody, R. W. *Methods in Enzymology*; Elsevier, 2004; Vol. *383*, pp 318−351.

(6) Nettels, D.; Müller-Späth, S.; Küster, F.; Hofmann, H.; Haenni, D.; Rüegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; et al. Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 20740−20745.

(7) Best, R. B.; Hummer, G. Optimized molecular dynamics force fields applied to the helix- coil transition of polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004−9015.

(8) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926.

(9) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665−9678.

(10) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.

(11) Zerze, G. H.; Best, R. B.; Mittal, J. Sequence-and temperature-dependent properties of unfolded and disordered proteins from atomistic simulations. *J. Phys. Chem. B* **2015**, *119*, 14622−14630.

(12) Best, R. B.; Zheng, W.; Mittal, J. Balanced Protein−Water Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. *J. Chem. Theory Comput.* **2014**, *10*, 5113−5124.

(13) Abascal, J. L.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.
(14) Neumann, M. Dielectric relaxation in water. Computer simulations with the TIP4P potential. *J. Chem. Phys.* **1986**, *85*, 1567−1580.
(15) van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. A systematic study of water models for molecular simulation: derivation of water models optimized for use with a reaction field. *J. Chem. Phys.* **1998**, *108*, 10220−10230.
(16) Höchtl, P.; Boresch, S.; Bitomsky, W.; Steinhauser, O. Rationalization of the dielectric properties of common three-site water models in terms of their force field parameters. *J. Chem. Phys.* **1998**, *109*, 4927−4937.
(17) Wu, Z.; Cui, Q.; Yethiraj, A. A new coarse-grained model for water: the importance of electrostatic interactions. *J. Phys. Chem. B* **2010**, *114*, 10524−10529.
(18) Braun, D.; Boresch, S.; Steinhauser, O. Transport and dielectric properties of water and the influence of coarse-graining: Comparing BMW, SPC/E, and TIP3P models. *J. Chem. Phys.* **2014**, *140*, 064107.
(19) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.
(20) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71.
(21) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **2015**, *119*, 5113−5123.
(22) Jurij, R.; Per, L. MOLSIM: A modular molecular simulation software. *J. Comput. Chem.* **2015**, *36*, 1259−1274.
(23) Raj, P. A.; Edgerton, M.; Levine, M. Salivary histatin 5: dependence of sequence, chain length, and helical conformation for candidacidal activity. *J. Biol. Chem.* **1990**, *265*, 3898−3905.
(24) Tsai, H.; Raj, P. A.; Bobek, L. A. Candidacidal activity of recombinant human salivary histatin-5 and variants. *Infect. Immun.* **1996**, *64*, 5000−5007.
(25) Das, R. K.; Ruff, K. M.; Pappu, R. V. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2015**, *32*, 102−112.
(26) Berendsen, H.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(27) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306−317.
(28) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.; Berendsen, H. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(29) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(30) Berendsen, H.; Van Gunsteren, W. Practical algorithms for dynamic simulations. *Molecular-Dynamics Simulation of Statistical-Mechanical Systems* **1986**, 43−65.
(31) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An *N* log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089.
(32) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
(33) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys. (Melville, NY, U. S.)* **1981**, *52*, 7182.
(34) Hess, B.; Bekker, H.; Berendsen, H.; Fraaije, J. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(35) Schrödinger, LLC. *PyMOL Molecular Graphics System*, version 1.2r1; 2009.
(36) Cragnell, C.; Durand, D.; Cabane, B.; Skepö, M. Coarse-grained modelling of the intrinsically disordered protein Histatin 5 in solution: Monte Carlo simulations in combination with SAXS. *Proteins: Struct., Funct., Genet.* **2016**, *84*, 777−791.
(37) Jephthah, S.; Henriques, J.; Cragnell, C.; Puri, S.; Edgerton, M.; Skepö, M. Structural Characterization of Histatin 5−Spermidine Conjugates: A Combined Experimental and Theoretical Study. *J. Chem. Inf. Model.* **2017**, *57*, 1330−1341.
(38) Cragnell, C.; Rieloff, E.; Skepö, M. Utilizing Coarse-Grained Modeling and Monte Carlo Simulations to Evaluate the Conformational Ensemble of Intrinsically Disordered Proteins and Regions. *J. Mol. Biol.* **2018**, *430*, 2478.
(39) Kabsch, W.; Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577−2637.
(40) Chebrek, R.; Leonard, S.; de Brevern, A. G.; Gelly, J.-C. PolyprOnline: polyproline helix II and secondary structure assignment database. *Database* **2014**, *2014*, No. bau102.
(41) Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols* **2007**, *1*, 2876.
(42) Wu, D.; Chen, A.; Johnson, C. An Improved Diffusion-Ordered Spectroscopy Experiment Incorporating Bipolar-Gradient Pulses. *J. Magn. Reson., Ser. A* **1995**, *115*, 260−264.
(43) Wilkins, D. K.; Grimshaw, S. B.; Receveur, V.; Dobson, C. M.; Jones, J. A.; Smith, L. J. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* **1999**, *38*, 16424−16431.
(44) Nygaard, M.; Kragelund, B. B.; Papaleo, E.; Lindorff-Larsen, K. An efficient method for estimating the hydrodynamic radius of disordered protein conformations. *Biophys. J.* **2017**, *113*, 550−557.
(45) Best, R. B.; Buchete, N.-V.; Hummer, G. Are current molecular dynamics force fields too helical? *Biophys. J.* **2008**, *95*, L07−L09.
(46) Henriques, J.; Cragnell, C.; Skepö, M. Molecular dynamics simulations of intrinsically disordered proteins: force field evaluation and comparison with experiment. *J. Chem. Theory Comput.* **2015**, *11*, 3420−3431.
(47) Henriques, J.; Skepö, M. Molecular dynamics simulations of intrinsically disordered proteins: on the accuracy of the TIP4P-D water model and the representativeness of protein disorder models. *J. Chem. Theory Comput.* **2016**, *12*, 3407−3415.
(48) Huang, J.; MacKerell, A. D. Force field development and simulations of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2018**, *48*, 40−48.
(49) Hicks, A.; Zhou, H.-X. Temperature-induced collapse of a disordered peptide observed by three sampling methods in molecular dynamics simulations. *J. Chem. Phys.* **2018**, *149*, 072313.
(50) Malmberg, C.; Maryott, A. Dielectric Constant of Water from 00 to 1000 C. *J. Res. Natl. Bur. Stand* **1956**, *56*, 2641.
(51) Owen, B. B.; Miller, R. C.; Milner, C. E.; Cogan, H. L. The Dielectric Constant of Water As a Function of Temperature and Pressure1, 2. *J. Phys. Chem.* **1961**, *65*, 2065−2070.
(52) Uematsu, M.; Frank, E. Static dielectric constant of water and steam. *J. Phys. Chem. Ref. Data* **1980**, *9*, 1291−1306.
(53) Fernández, D. P.; Goodwin, A.; Lemmon, E. W.; Levelt Sengers, J.; Williams, R. A formulation for the static permittivity of water and steam at temperatures from 238 to 873 K at pressures up to 1200 MPa, including derivatives and Debye−Hückel coefficients. *J. Phys. Chem. Ref. Data* **1997**, *26*, 1125−1166.
(54) Kurut, A.; Henriques, J.; Forsman, J.; Skepö, M.; Lund, M. Role of histidine for charge regulation of unstructured peptides at interfaces and in bulk. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 657−667.
(55) Chemes, L. B.; Alonso, L. G.; Noval, M. G.; de Prat-Gay, G. In *Intrinsically Disordered Protein Analysis*; Uversky, V. N., Dunker, A. K., Eds.; Humana Press: Totowa, NJ, 2012; Vol. *895*; pp 387−404.

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672−2683

(56) Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V. N. Introducing protein intrinsic disorder. *Chem. Rev.* **2014**, *114*, 6561−6588.

DOI: 10.1021/acs.jctc.8b01281
J. Chem. Theory Comput. 2019, 15, 2672−2683